Illinois Journal of Mathematics
Volume 45, Number 2, Summer 2001, Pages 559–580
S 0019-2082

PAIR CORRELATIONS AND U-STATISTICS FOR INDEPENDENT AND WEAKLY DEPENDENT RANDOM VARIABLES

ISTVÁN BERKES, WALTER PHILIPP, AND ROBERT TICHY

Abstract. We prove a Glivenko-Cantelli type strong law of large numbers for the pair correlation of independent random variables. Except for a few powers of logarithms the results obtained are sharp. Similar estimates hold for the pair correlation of lacunary sequences $\{n_k\omega\}$ mod 1.

1. Introduction

Pair correlations have been studied during the last seventy years in various forms and disguises in mathematics, statistics, fluid mechanics, electrical engineering and various branches of physics. Let $\{X_j : 1 \leq j \leq N\}$ be real numbers or vectors, or random variables. In some form or other pair correlations are measuring the average closeness of the $N(N-1)$ pairs $(X_i, X_j), 1 \leq i \neq j \leq N$. Here "closeness" is to be understood in very general sense, not just with respect to the Euclidean distance. With this in mind the concept of pair correlations can be traced back at least to the 1930s when, under the name "cluster integrals", it has started to play an important role in fluid mechanics. See, e.g., Green [12, p. 91] and the literature quoted there. In Killingbeck and Cole [16, p. 589], the term "pair correlation functions" has a slightly different meaning. For applications in electrical engineering we refer to Hess and Sah [13]. In statistics the concept was investigated, among others, by Eberl and Hafner [10], Silverman [24], Horvath [14] and Eastwood and Horvath [9]. Montgomery [19] investigated the pair correlation of the zeros of the Riemann zeta function under the assumption of the Riemann hypothesis, a subject taken up very recently in a profound paper of Rudnick and Sarnak [21]. The connections between uniform distribution mod 1, pair correlations,

Received February 8, 2000; received in final form July 10, 2000.
2000 *Mathematics Subject Classification.* 11K06, 60F15.
This work was partially supported by the Austrian Science Foundation project S 8308-MAT. The paper was completed when the third author visited the Mathematics Department of the University of Illinois at Urbana-Champaign.

and the spacings between the energy levels of harmonic oscillators were recently investigated by Rudnick and Sarnak [22]. This, as well as a very recent paper by Rudnick and Zaharescu [23] provided the impetus for the present paper.

Several of the papers mentioned above deal with deterministic, i.e., non-random sequences $\{X_j, j \geq 1\}$. For random sequences $\{X_j, j \geq 1\}$, i.e., for sequences of random variables or random vectors the concept of pair correlations as well as the closely related concept of correlation integrals all fall under the umbrella of empirical processes of U-statistic structure or U-processes. We introduce these objects on a level more general than needed for presentation of our results, yet on a level general enough to demonstrate the connections between these concepts.

Let $\{X_j, j \geq 1\}$ be a sequence of identically distributed p-dimensional random vectors with common distribution function F. Let $\mathcal{H}$ be a class of kernel functions $h : \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}$. Usually it is assumed that h is symmetric in the first two variables, i.e., $h(x, y, t) = h(y, x, t)$ for all $x, y \in \mathbb{R}^p$ and $t \in \mathbb{R}$. However, in some of our results (Propositions 1 and 2) we also need non-symmetric kernels. More generally, the X_j's can be Banach space valued, the parameter t can be replaced by classes of functions, such as V-C classes, and the multivariate situation is a straight-forward extension of the present case (m=2). The U-process indexed by the class $\mathcal{H}$ can be defined as

$$U_N(t) = U_N(h; t) := \sum_{1 \leq i < j \leq N} h(X_i, X_j, t) \ , \quad N \geq 1 \ , h \in \mathcal{H} \ . \tag{1.1}$$

In the independent case $\frac{2}{N(N-1)} U_N(t)$ is an unbiased estimator for

$$U(t) = \int_{\mathbb{R}^{2p}} h(x, y, t) dF(x) dF(y).$$

In the present paper we shall make two major assumptions. First, we restrict ourselves to the case of random variables X_j all of which have uniform distribution over $[0, 1]$, i.e.,

$$\mathbb{P}(X_j \leq x) = x \ , \quad 0 \leq x \leq 1, \ j \geq 1 \ . \tag{1.2}$$

Stationarity of the sequence $\{X_j, j \geq 1\}$ is not assumed in the case of dependent random variables. Second, we shall restrict ourselves to the following two kernels:

$$h(x, y, t) = \mathbf{1}\left(|x - y| \leq t\right) \tag{1.3}$$

and

$$\begin{aligned} h(x, y, t) &= \mathbf{1}\left(0 \leq \{x\} - \{y\} \leq t \mod 1\right) \\ &= \mathbf{1}\left(\{x - y\} \leq t\right) = \mathbf{1}_{[0,t]}(x - y). \end{aligned} \tag{1.4}$$

Here $\mathbf{1}(A)$ denotes the indicator of the set A and $\mathbf{1}_A(\cdot)$ also denotes the indicator of the set A, but extended with period 1; $\{x\} = x - [x]$ denotes the fractional part of x. In (1.3) $|\cdot|$ is the Euclidean distance in $\mathbb{R}^1$, whereas in (1.4) $\{x\}$ and $\{y\}$ are considered points on the one-dimensional torus group.

For the kernel h defined in (1.3) (and for a general distribution function F) the corresponding U-process

$$U_N(t) := \sum_{1\le i<j\le N} \mathbf{1}\left(|X_j - X_i| \le t\right) \tag{1.5}$$

divided by $\frac{N(N-1)}{2}$ is called the empirical or sample correlation integral. If, in addition, $t = t_N \to 0$ the corresponding limit theorems often appear under the heading "limit theorems for short distances", as well as under the heading "pair correlations". (See, e.g., Eastwood and Horvath [9], Horvath [14], Rudnick and Sarnak [22] and Rudnick and Zaharescu [23].) For stationary ergodic sequences $\{X_j, j \ge 1\}$, with a general common distribution function F, and for a very general class of kernels h, Borovkova, Burton and Dehling [2] established a Glivenko-Cantelli type theorem proving almost sure uniform convergence of $\binom{N}{2}^{-1} U_N(t)$ to $U(t)$. For strictly stationary absolutely regular sequences $\{X_j, j \ge 1\}$ they proved a weak invariance principle with the approximating process being a multiple stochastic integral with integrators being specified Kiefer processes. This extends an earlier result of Dehling, Denker, and Philipp [7] who obtained this result for sequences of independent, identically distributed random variables.

The purpose of the present paper is to prove almost sure uniform estimates of $U_N(t)$, given in (1.5), over short intervals. We shall provide error terms which, except for powers of logarithm, are sharp.

We now state our results in detail. In case that the kernel h is given by (1.3) the corresponding U-process is given by (1.5). We set for $0 \le \alpha \le 2$

$$\begin{aligned} \Gamma_N(\alpha) : &= \sup\left(\left| U_N(t) - \binom{N}{2} \cdot (2t - t^2) \right| : t \le N^{-\alpha} \right) \\ &= \sup\left(\Big| \sum_{1\le i<j\le N} \left(\mathbf{1}\left(|X_j - X_i| \le t\right) - (2t - t^2)\right) \Big| : t \le N^{-\alpha} \right) . \end{aligned} \tag{1.6}$$

The first result is a strong law of large numbers for $\Gamma_N(\alpha)$, in the case that $\{X_j, j \ge 1\}$ is a sequence of independent random variables with uniform distribution over $[0,1]$. This result can be reinterpreted as a Glivenko-Cantelli type strong law of large numbers over short intervals with a reasonably sharp error estimate.

THEOREM 1. *If $\{X_j, j \ge 1\}$ is a sequence of independent random variables with uniform distribution (1.2), then*

$$\Gamma_N(\alpha) \ll \left(N^{1-\frac{1}{2}\alpha} + N^{\frac{3-3\alpha}{2}} \right) (\log N)^{9/2}$$

with probability 1.

Aside from the power of $\log N$, this estimate is sharp.

THEOREM 2. *Let $0 \le \alpha \le 2$. Then with probability 1*

$$\limsup_{N\to\infty} (N^{1-\frac{1}{2}\alpha} + N^{\frac{3-3\alpha}{2}})^{-1}\Gamma_N(\alpha) \ge c > 0 \ ,$$

for some constant c.

We now treat the case (1.4) and denote the corresponding U-process by

$$V_N = V_N(L) := \sum_{1\le i<j\le N} \mathbf{1}_L\left(X_j - X_i\right) \ . \tag{1.7}$$

For $0 \le \alpha \le 2$ we set

$$\Delta_N(\alpha) := \sup\left\{\left|V_N(L) - \binom{N}{2}|L|\right| : L \subseteq [0, N^{-\alpha}]\right\} \ . \tag{1.8}$$

Here $|L|$ denotes the length of the interval $L \subseteq [0,1]$.

PROPOSITION 1. *If $\{X_j, j \ge 1\}$ is a sequence of independent random variables with uniform distribution (1.2) then*

$$\Delta_N(\alpha) \ll N^{1-\frac{1}{2}\alpha}(\log N)^{9/2}$$

with probability 1.

REMARK 1. In fact, both Proposition 1 and Theorem 1 remain valid for $\alpha > 2$.

While skimming through the proof of Proposition 1 the connoisseur undoubtedly will recognize that in this case a bounded Chung-Smirnov law of the iterated logarithm holds in the form

$$\Delta_N(\alpha) \ll N^{1-\frac{1}{2}\alpha} \log\log N \quad \text{a.s.} \tag{1.9}$$

We will give a sketch of the proof of (1.9) at the end of Section 5, after the underlying martingale structure has been reasonably well developed.

Again, except for powers of logarithm our estimate is sharp.

PROPOSITION 2. *Let $0 \le \alpha \le 2$. Then with probability 1*

$$\limsup_{N\to\infty} N^{-(1-\frac{1}{2}\alpha)}\Delta_N(\alpha) \ge c > 0$$

for some constant c.

REMARK 2. It is perhaps interesting to compare the estimates for Δ_N and Γ_N. The additional term $N^{\frac{3-3\alpha}{2}}$ stems from the fact that the kernel corresponding to Δ_N is periodic with period 1 whereas the other kernel is not.

We now shift our focus to lacunary sequences $\{n_j\omega\} \mod 1$. In a recent paper, Rudnick and Zaharescu [23] proved the following interesting result: Let $\{n_j, j \geq 1\}$ be a sequence of integers satisfying a Hadamard gap condition, i.e.,

$$n_{j+1}/n_j \geq q > 1 \tag{1.10}$$

for some $q > 1$. For $0 \leq \omega < 1$ set

$$X_j = n_j\omega \mod 1 . \tag{1.11}$$

Now $\{X_j, j \geq 1\}$ is a sequence of (weakly dependent) random variables, defined on $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1)B, \mathbb{P})$ where B are the Lebesgue sets and $\mathbb{P}$ is the Lebesgue measure. As is easy to see each X_j has uniform distribution over $[0,1)$. In this situation Rudnick and Zaharescu [23] proved that

$$\frac{1}{N}U_N(t/N) \to t \quad \text{with probability } 1 ,$$

i.e., for all $0 \leq \omega < 1$ except on a set of Lebesgue measure 0. Here $U_N(t)$ is defined as in (1.5).

One of the purposes of the present paper is to improve upon this result in several ways. First, we assume that $\{n_j, j \geq 1\}$ satisfies an Erdős gap condition, i.e.,

$$n_{j+1}/n_j \geq 1 + cj^{-\rho} \ , \ j \geq 1 \tag{1.12}$$

for some $c > 0$ and $0 \leq \rho < 1$. Notice that if we set $\rho = 0$ then condition (1.12) reduces to the Hadamard gap condition (1.10). Second, we allow for longer intervals. Third, we will prove uniform convergence over all subintervals, not just those that are symmetric about 0. Finally, we give an error term on the speed of convergence (which at least in the case of independent random variables is sharp, aside from the power of $\log N$, according to Theorem 2). Our results are as follows.

THEOREM 3. *Let $\{n_j, j \geq 1\}$ be a sequence of integers satisfying an Erdős gap condition (1.12). Let X_j be defined by (1.11) and set $\Gamma_N(\alpha) = \Gamma_N(\alpha, \rho)$ as in (1.6). Then with probability 1*

$$\Gamma_N(\alpha) \ll \left(N^{1-\frac{1}{2}\alpha+\frac{3}{2}\rho} + N^{2-2\alpha}\right)(\log N)^{11/2} \ .$$

PROPOSITION 3. *Let $\{n_j, j \geq 1\}$ be a sequence of integers satisfying an Erdős gap condition (1.12). Let X_j be defined by (1.11) and set $\Delta_N(\alpha) = \Delta_N(\alpha, \rho)$ as in (1.8). Then with probability* 1

$$\Delta_N(\alpha) \ll N^{1-\frac{1}{2}\alpha+\frac{3}{2}\rho}(\log N)^{11/2} \ .$$

In case that $\alpha \leq 3\rho$ we obtain a much stronger result.

PROPOSITION 4. *Let $\{n_j, j \geq 1\}$ be any increasing sequence of integers. Let X_j be defined by (1.11). Then with probability 1*

$$\Delta_N(0) \ll N(\log N)^{2+\varepsilon} \ , \quad \varepsilon > 0 \ .$$

Proposition 3 and similar results for different sequences $\{n_j, j \geq 1\}$ will follow from the following theorem, which is similar in spirit as Theorem 1 of Rudnick and Zaharescu [23]. First, we introduce some notation. Let $\{n_j, j \geq 1\}$ be an increasing sequence of integers. Let $1 \leq Q \leq H$ be integers and let $d > 0$ and $\sigma \geq 0$. Consider the system of Diophantine equations

$$\left\{ \begin{array}{ll} a(n_j - n_i) = b(n_l - n_k) & (H \leq j < H+Q, \ 1 \leq i < j, \\ & H \leq l < H+Q, \ 1 \leq k < l) \\ 1 - dH^{-\sigma} \leq \left|\frac{a}{b}\right| \leq 1 + dH^{-\sigma} & (a, b \in \mathbb{Z}) \ . \end{array} \right. \tag{1.13}$$

We call a solution $(a, b, n_i, n_j, n_k, n_l)$ trivial if both $j = l$ and $i = k$. For fixed $a \in \mathbb{Z}$ there are at most $2HQ$ trivial solutions, even if $d = \infty$.

PROPOSITION 5. *Let $\{n_j, j \geq 1\}$ be a sequence of integers. Suppose there exist constants $c > 0, C > 0, \sigma > 0, \gamma \geq 0, \tau \geq 1$ with the following property. For each pair (Q, H) with $1 \leq Q \leq H$ and for each fixed $a \in \mathbb{Z}$ with*

$$0 < |a| \leq H^{20} \tag{1.14}$$

the system (1.13) of Diophantine equations has at most $CH^\tau Q(\log H)^\gamma$ solutions. Then with probability 1

$$\Delta_N(\alpha) \ll N^{\frac{1}{2}(1+\sigma+\tau-\alpha)}(\log N)^{\frac{1}{2}(\gamma+9)} \ .$$

2. Proof of Proposition 5

Recall that $\mathbf{1}_L$ has period 1. We write

$$S = S(H, Q; L; \omega) = \sum_{H \leq j < H+Q} \ \sum_{1 \leq i < j} (\mathbf{1}_L((n_j - n_i)\omega) - |L|) \ .$$

LEMMA 1. *Under the hypotheses of Proposition 5 we have*

$$\mathbb{E}S^2 \ll |L| H^{\sigma+\tau} Q(\log H)^{\gamma+1} + H^{-16} \ ,$$

where the constant implied by $\ll$ only depends on C.

Proof. We set

$$\begin{aligned} f(\omega) &= \mathbf{1}_L(\omega) - |L| = \sum_{a \neq 0} c_a e^{2\pi\sqrt{-1}\, a\omega} \ , \\ f_H^*(\omega) &= \sum_{0 < |a| \leq H^{20}} c_a e^{2\pi\sqrt{-1}\, a\omega} \end{aligned}$$

and

$$f_H^{**}(\omega) = f(\omega) - f_H^*(\omega).$$

Clearly $|c_a| \le |a|^{-1}$ and thus

$$||f_H^{**}||_2^2 \le \sum_{|a|>H^{20}} |a|^{-2} \le H^{-20} \ . \tag{2.1}$$

Now set

$$S_H^*(\omega) := \sum_{H\le j<Q+H} \sum_{1\le i<j} f_H^*((n_j-n_i)\omega) \ .$$

By Minkowski's inequality and by (2.1)

$$||S-S_H^*||_2 \le Q(Q+H)H^{-10} \ll H^{-8} \ . \tag{2.2}$$

Thus we need to estimate $||S_H^*||_2$. We have for some constant θ

$$||S_H^*||_2 \le \sum_{0\le u\le \theta H^\sigma \log H} \left|\left| \sum_a c_a \sum_{H\le j<H+Q} \sum_{1\le i<j} e^{2\pi\sqrt{-1}a(n_j-n_i)\omega} \right|\right|_2 \ ,$$

where $\sum_a$ is extended over all a with $(1+dH^{-\sigma})^u \le |a| \le (1+dH^{-\sigma})^{u+1}$.
Thus, since $|xy| \le \frac{1}{2}(x^2+y^2)$, we obtain

$$\begin{aligned} ||S_H^*||_2 \le & \sum_{0\le u\le \theta H^\sigma \log H} \Bigg(\sum_{a,b} c_a\bar{c}_b \sum_{H\le j,l<H+Q} \\ & \qquad \sum_{\substack{1\le i<j\\ 1\le k<l}} \mathbf{1}\left(a(n_j-n_i)=b(n_l-n_k)\right) \Bigg)^{1/2} \\ \le & \sum_{0\le u\le \theta H^\sigma \log H} \Bigg(\sum_a |c_a|^2 \sum_b \sum_{H\le j,l<H+Q} \\ & \qquad \sum_{\substack{1\le i<j\\ 1\le k<l}} \mathbf{1}\left(a(n_j-n_i)=b(n_l-n_k)\right) \Bigg)^{1/2} \ . \end{aligned}$$

The inner most triple sum does not exceed the number of solutions of the system (1.13). Thus using the main hypothesis of Proposition 5 we obtain by Cauchy's inequality and Parseval's identity

$$\begin{aligned} ||S_H^*||_2 &\le (CH^\tau Q(\log H)^\gamma)^{1/2} \sum_{0\le u\le \theta H^\sigma \log H} \left(\sum_a |c_a|^2 \right)^{1/2} \\ &\ll (H^\tau Q(\log H)^\gamma)^{1/2} (H^\sigma \log H)^{1/2} \cdot |L|^{1/2} \ . \end{aligned}$$

The lemma follows now from (2.2). $\square$

Lemma 2. *Fix $0 \le \alpha \le 2$, and set*

$$\Delta = \Delta(H,Q;\alpha;\omega) = \sup\left\{|S(H,Q;L;\omega| : L \subset [0,H^{-\alpha}]\right\} \ .$$

Then for each $R \ge 1$

$$\mathbb{P}(\Delta \ge R) \ll R^{-2} Q H^{\sigma+\tau-\alpha} (\log H)^{\gamma+4} \ .$$

Proof. Denote a typical interval L by $L = H^{-\alpha}[0,s], 0 \le s \le 1$. We write s in its binary expansion

$$s = \sum_{m \ge 1} \varepsilon_m 2^{-m} \qquad \varepsilon_m = 0,1 \ , \tag{2.3}$$

and we set

$$M = [4 \log H] \ . \tag{2.4}$$

Then

$$s = \sum_{m=1}^{M} \varepsilon_m 2^{-m} + \theta 2^{-M} \ , \tag{2.5}$$

where $0 \le \theta \le 1$. For $0 \le u < v \le 1$ we set

$$\begin{aligned} Z(u,v) &= Z(u,v;H,Q) \\ &= \sum_{H \le j < H+Q} \sum_{1 \le i < j} \left(\mathbf{1}_{H^{-\alpha}[u,v)}\left((n_j - n_i)\omega\right) - (v-u)H^{-\alpha}\right) \ . \end{aligned}$$

Then for each $0 \le s \le 1$

$$|Z(0,s)| \le \sum_{m=1}^{M} |Z(d_m 2^{-m}, (d_m+1)2^{-m})| + 2Q(H+Q)2^{-M} H^{-\alpha} \ , \tag{2.6}$$

where the d_m's are integers with $0 \le d_m < 2^m$ for $1 \le m \le M$. This is a special case of the familiar chaining argument in empirical process theory, dating back at least 50 years; see, e.g., Cassels [4] or Chung [5]. Hence

$$\begin{aligned} \mathbb{P}\left(\sup_{0 \le s \le 1} |Z(0,s)| \ge 2R\right) &\le \sum_{m=1}^{M} \mathbb{P}\left(\max_{0 \le d < 2^m} |Z(d2^{-m},(d+1)2^{-m})| \ge R/M\right) \\ &\le \sum_{m=1}^{M} \frac{M^2}{R^2} 2^m (2^{-m} H^{-\alpha} Q H^{\sigma+\tau} (\log H)^{\gamma+1} + H^{-16}) \\ &\ll R^{-2} Q H^{\sigma+\tau-\alpha} (\log H)^{\gamma+4} \ . \end{aligned} \tag{2.7}$$

□

We now can finish the proof of Proposition 5. Fix $r \ge 0$. We shall estimate

$$\max_{0 \le Q < 2^r} \Delta(2^r, Q; \alpha) \ .$$

Write Q in binary expansion. Then for each $0 \le Q < 2^r$

$$\Delta(2^r, Q; \alpha) \le \sum_{l=1}^{r} \Delta(2^r + m_l 2^{l-1}, 2^{l-1}; \alpha) \ ,$$

where $0 \le m_l \le 2^{r-l}$. This argument goes back to at least Gál and Koksma [11]. Hence, setting $\beta = \frac{1}{2}(1 + \sigma + \tau - \alpha)$, we have

$$\begin{aligned} &\mathbb{P}\left(\max_{0 \le Q < 2^r} \Delta(2^r, Q; \alpha) \ge 2^{r\beta} r^{\frac{\gamma+9}{2}} \right) \\ &\qquad \le \sum_{l=1}^{r} \mathbb{P}\left(\max_{0 \le m \le 2^{r-l}} \Delta(2^r + m 2^{l-1}, 2^{l-1} \ \alpha) \ge 2^{r\beta} r^{\frac{\gamma+7}{2}} \right) . \end{aligned}$$

We estimate these probabilities using Lemma 2 with $R = 2^{r\beta} r^{\frac{1}{2}(\gamma+7)}$, $H = 2^r + m 2^{l-1} \le 2^{r+1}$ and $Q = 2^{l-1}$. Then the probability in question is

$$\ll \sum_{l \le r} 2^{r-l} 2^{-2r\beta} r^{-\gamma-7} 2^l 2^{r(\sigma+\tau-\alpha)} r^{\gamma+4} \ll r^{-2} \ .$$

Replacing the last exponent -2 by $-1-\varepsilon$ will give a smaller power of $\log N$. Thus by the Borel-Cantelli lemma we have with probability 1

$$\max_{0 \le Q \le 2^r} \Delta(2^r, Q; \alpha) \ll 2^{r\beta} r^{\frac{1}{2}(\gamma+9)} \ .$$

If N is given, define n by $2^{n-1} \le N < 2^n$. Then with probability 1

$$\Delta_N(\alpha) \ll \sum_{r<n} \max_{0 \le Q \le 2^r} \Delta(2^r, Q; \alpha) \ll \sum_{r<n} 2^{r\beta} r^{\frac{1}{2}(\gamma+9)} \ll N^\beta (\log N)^{\frac{1}{2}(\gamma+9)} \ ,$$

which yields Proposition 5.

Notice that in the last estimate we used the fact that as a function of r the suprema over intervals contained in $[0, 2^{-r\alpha}]$ are non-increasing.

3. Proof of Proposition 3

The first lemma is well-known.

Lemma 3. *Suppose that $\{n_j, j \ge 1\}$ satisfies (1.12). Let $A < B$. Then the number of n_j's with $j \le N$ and $A \le n_j \le B$ is $\ll N^\rho \log(B/A)$, where the constant implied by $\ll$ only depends on c and ρ.*

Proof. Let n_k be the largest member of the sequence not exceeding B and let n_j be the smallest member $\ge A$, so that

$$A \le n_j < n_k \le B \ .$$

Since

$$B/A \ge n_k/n_j \ge (1 + c(k-1)^{-\rho}) \dots (1 + cj^{-\rho}) \ge (1 + ck^{-\rho})^{k-j}$$

we have

$$k - j \le \log(B/A)(\log(1 + ck^{-\rho}))^{-1} \ll N^{\rho} \log(B/A). \qquad \square$$

Lemma 4. *Suppose that $\{n_j, j \ge 1\}$ satisfies (1.12). Let $1 \le Q \le H$ be given and fix $a \in \mathbb{Z}$ with*

$$0 < |a| \le H^{20} \ .$$

Set $\sigma = \rho$ and $d = \frac{c}{4}$. Then the system (1.13) of Diophantine equations has at most $CH^{1+2\rho}Q(\log H)^2$ solutions. Here $C > 0$ depends only on c and ρ.

Proof. Assume that $(a, b, n_i, n_j, n_k, n_l)$ is a solution of the system (1.13). We treat the case $a > 0, b > 0$ only. The other three cases can be treated similarly. Let $H \ge H_0$, where $H_0 = H_0(c, \rho)$ depends only on c and ρ. Then

$$\frac{1}{3} cn_j H^{-\rho} \le n_j - n_i < n_j$$

and

$$\frac{1}{3} cn_l H^{-\rho} \le n_l - n_k < n_l \ .$$

Thus by (1.13)

$$\frac{n_j}{n_l} \cdot \frac{1}{3} cH^{-\rho} \le \frac{n_j - n_i}{n_l - n_k} = \frac{b}{a} \le \frac{n_j}{n_l} \cdot \frac{3}{c} H^{\rho} \ ,$$

and hence using the bounds on a/b we obtain for $H \ge H_0$

$$\frac{1}{6} cH^{-\rho} \le \frac{n_j}{n_l} \le \frac{6}{c} H^{\rho} \ .$$

Hence by Lemma 3 with $A = \frac{c}{6} H^{-\rho} n_l$ and $B = \frac{6}{c} H^{\rho} n_l$ we conclude that at most $4\rho H^{\rho} Q \log H$ pairs (j, l) possibly can qualify.

Fix such a pair and assume $j > l$. The case $j < l$ can be treated similarly, and the case $j = l$ will be treated separately below. Set

$$\nu := an_j - bn_l \ .$$

Then, as $j > l$, we obtain from (1.13)

$$\nu = an_j \left(1 - \frac{b}{a} \frac{n_l}{n_j}\right) \ge \frac{1}{4} \cdot an_j \cdot cH^{-\rho} \ .$$

But we also have $\nu = an_i - bn_k < an_i$, and so

$$n_j \cdot \frac{1}{4} cH^{-\rho} \le \frac{\nu}{a} \le n_i \le n_j(1 + cj^{-\rho})^{i-j} \le n_j(1 + c(2H)^{-\rho})^{i-j} \ . \tag{3.1}$$

Consequently,

$$(j - i)cH^{-\rho} \le 2\rho \log H \ .$$

This implies $j - i \le CH^{\rho} \log H$. Since there are at most $2H$ choices for k and since finally b is determined by (1.13) we see that there are at most

$CH^{1+2\rho}\log^2 H$ nontrivial solutions of (1.13). Now, assume $j=l$. Then (1.13) reduces to

$$(a-b)n_j = an_i - bn_k \qquad \text{with} \quad 1\le i,k<j\ . \tag{3.2}$$

Assume $b<a$. The case $b>a$ can be treated similarly, and $b=a$ implies $n_i=n_k$, which is the trivial case. Now (3.1) and (3.2) imply

$$n_j \le (a-b)n_j = an_i - bn_k < an_i < H^{20} n_j (1+c(2H)^{-\rho})^{i-j}\ ,$$

and so $j-i\le 40c^{-1}H^\rho \log H$. Again, there are at most $2H$ possibilities for k and b is determined by (1.13). □

Applying Proposition 5 with $\sigma=\rho, \tau=1+2\rho$ and $\gamma=2$, we obtain Proposition 3.

4. Proof of Proposition 4

By the Erdős-Turán inequality (see, e.g., Drmota and Tichy [8, p. 15] or Kuipers and Niederreiter [18, p. 112]) we have for each $M>0$

$$\begin{aligned}\Delta_N(0) &\le 6N^2M^{-1} + 6\sum_{h=1}^{M}\frac{1}{h}\left|\sum_{1\le j\ne k\le N} e^{2\pi i(n_j-n_k)h\omega}\right| \\ &\le 6N^2M^{-1} + 6\sum_{h=1}^{M}\frac{1}{h}\left(\left|\sum_{1\le j\le N} e^{2\pi i n_j h\omega}\right|^2 + N\right)\ . \end{aligned}\tag{4.1}$$

A straight-forward adaption of Baker's argument (see [1, pp. 37–38]) yields the result. The details are as follows: By Hunt's maximal inequality ([15]; see also Baker [1, relation (10)]),

$$\mathbb{E}\left(\max_{\frac12 n\le r\le n}\left|\sum_{j\le r} e^{2\pi i n_j h\omega}\right|^2\right) \le Cn\ ,$$

where C is a numerical constant. Thus by (4.1) with $M=N(=m)$ we have for $k\ge 1$

$$\begin{aligned}&\mathbb{E}\left(\max_{2^{k-1}\le m<2^k}\Delta_m(0)\right) \\ &\qquad \le 6\cdot 2^k + 6\sum_{h\le 2^k}\frac{1}{h}\left(\mathbb{E}\max_{2^{k-1}\le r<2^k}\left|\sum_{j\le r} e^{2\pi i n_j h\omega}\right|^2 + 2^k\right) \ll 2^k\cdot k\ .\end{aligned}$$

Thus by Markov's inequality and the Borel-Cantelli Lemma we have with probability 1

$$\max_{2^{k-1}\le m<2^k}\Delta_m(0) \ll 2^k k^{2+\varepsilon}\ ,$$

which finally yields Proposition 4.

5. Proofs of Propositions 1 and 2 and Theorem 2

The proof of Proposition 1 is an easy modification of the proof of Proposition 5. For $1 \le Q \le H$ set

$$S = S(H, Q; L; \omega) = \sum_{H \le j < H+Q} \sum_{1 \le i < j} \left(\mathbf{1}_L(X_j - X_i) - |L| \right) .$$

We first show that the conclusion of Lemma 1 remains valid.

LEMMA 5. *We have*

$$\mathbb{E} S^2 \ll |L| H Q \log H + H^{-16} .$$

Proof. We define f, f_H^*, f_H^{**} and, with the obvious modification, S_H^* as in the proof of Lemma 1. Then (2.2) remains valid. To estimate $||S_H^*||$ we proceed as in the proof of Lemma 1. We observe that because of independence the expectation of each term in the relevant sum

$$\sum_b \sum_{H \le j, l < H+Q} \sum_{\substack{1 \le i < j \\ 1 \le k < l}} e^{2\pi\sqrt{-1}(a(X_j - X_i) - b(X_l - X_k))}$$

vanishes unless $j = l, i = k$ and $b = a$. Thus for fixed a this sum reduces to $Q(H+Q)$. This yields the conclusion of the lemma. □

The remainder of the proof of Proposition 1 is identical to that of Proposition 5, in the special case $\sigma = 0, \tau = 1$ and $\gamma = 0$.

For the proof of Theorem 2 it is enough to show that there is a positive constant c such that for all $N \ge 1$

$$\mathbb{P}\left(\left| \sum_{1 \le i < j \le N} \left(\mathbf{1}\left(|X_i - X_j| \le N^{-\alpha} \right) - (2N^{-\alpha} - N^{-2\alpha}) \right) \right| \ge c(N^{1-\frac{1}{2}\alpha} + N^{(3-3\alpha)/2}) \right) \ge c, \tag{5.1}$$

since then

$$\mathbb{P}\left(\Gamma_N(\alpha) \ge c(N^{1-\frac{1}{2}\alpha} + N^{(3-3\alpha)/2}) \right) \ge c , \tag{5.2}$$

and thus by the Kolmogorov zero-one law the right-hand side of (5.2) can be replaced by 1. But relation (5.1) follows from Horvath [14, Theorem 2] with $c(N) = N^{-\alpha}$.

For the proof of Proposition 2 again it is enough to show the existence of a positive constant c such that for all $N \geq 1$

$$\mathbb{P}\left(\left|\sum_{1\leq i<j\leq N}\left(\mathbf{1}_{[0,N^{-\alpha}]}\left(X_j-X_i\right)-N^{-\alpha}\right)\right|\geq cN^{1-\frac{1}{2}\alpha}\right)\geq c\ . \tag{5.3}$$

We need two well-known facts from analysis and martingale theory. The first is the Paley-Zygmund theorem.

Theorem A. *Let X be a non-negative random variable with finite second moment. Suppose that*

$$||X||_2 \leq \frac{1}{a}\,\mathbb{E}X$$

for some $0 \leq a \leq 1$. Let $0 \leq b \leq a$. Then

$$\mathbb{P}(X \geq b||X||_2) \geq (b-a)^2\ .$$

The next result is a special case of Burkholder's square function inequality (see [3, Theorem 3.1]).

Theorem B. *Let $\{Y_n, \mathcal{F}_n, n \geq 1\}$ be a real-valued martingale. Then for every p with $1 < p < \infty$ there are constants $c_p > 0$ and C_p such that*

$$\begin{aligned} c_p\mathbb{E}\left\{\sup_{n\geq 1}|Y_n|^p\right\} &\leq \mathbb{E}\left\{\left(Y_1^2+\sum_{n\geq 1}(Y_{n+1}-Y_n)^2\right)^{\frac{1}{2}p}\right\} \\ &\leq C_p\mathbb{E}\left\{\sup_{n\geq 1}|Y_n|^p\right\}\ . \end{aligned}$$

Let $L = [0, N^{-\alpha}]$ and write $g(z) = \mathbf{1}_L(z) - |L|$.

Lemma 6. *Let $\mathcal{F}_k := \sigma(X_1, \ldots, X_k), k \geq 1$ be the natural filtration of the sequence $\{X_j, j \geq 1\}$. Let*

$$v_k := \sum_{1\leq j<k} g(X_k - X_j)\ .$$

Then $\{v_k, \mathcal{F}_k, k \geq 1\}$ is a martingale difference sequence and $\{V_k^*, \mathcal{F}_k, k \geq 1\}$ is a martingale. Here $V_k^* = V_N - \mathbb{E}V_N$, where V_N is defined in (1.7).

Proof. Since $\mathbf{1}_L$ has period 1, the conditional expectation

$$\begin{aligned}\mathbb{E}(v_k|\mathcal{F}_{k-1}) &= \sum_{1\le j<k} \mathbb{E}\{g(X_k - X_j)|\mathcal{F}_{k-1}\} \\ &= \sum_{1\le j<k} \int_0^1 g(u - X_j)du \\ &= \sum_{1\le j<k} \int_0^1 (\mathbf{1}_L(u) - |L|)du = 0 \ ,\end{aligned}$$

which completes the proof of the lemma. □

Lemma 7. *We have for $N \ge N_0(\alpha)$*

$$\mathbb{E}V_N^{*2} \ge \frac{1}{4} N^{\,2-\alpha} \ .$$

Proof. Since the sequence $\{v_k, \mathcal{F}_k, k \ge 1\}$ is a martingale difference sequence, the random variables v_k are orthogonal. Thus

$$\mathbb{E}V_N^{*2} = \sum_{k\le N} \mathbb{E}v_k^2 \ .$$

Now

$$\begin{aligned}\mathbb{E}(v_k^2|\mathcal{F}_{k-1}) &= \mathbb{E}\left\{\left(\sum_{1\le j<k} g(X_k - X_j)\right)^2 |\mathcal{F}_{k-1}\right\} \\ &= \int_0^1 \left(\sum_{1\le j<k} g(u - X_j)\right)^2 du \\ &= \sum_{1\le i,j<k} \int_0^1 g(u - X_j)g(u - X_i)du \ .\end{aligned}$$

For $i \ne j$ the expectation of the integrals vanish by Fubini's theorem and independence. Thus by periodicity

$$\begin{aligned}\mathbb{E}v_k^2 &= \mathbb{E}(\mathbb{E}(v_k^2|\mathcal{F}_{k-1})) = \sum_{1\le j<k} \mathbb{E} \int_0^1 g^2(u - X_j)du \\ &= (k-1)\mathbb{E}\left(\int_0^1 \mathbf{1}_L(u)du - |L|^2\right) \\ &= (k-1)(N^{-\alpha} - N^{-2\alpha}) \ ,\end{aligned}$$

which yields the lemma. □

The next step is to estimate $\mathbb{E}V_N^{*4}$ and to apply Theorem A. For this we need the following result.

LEMMA 8. *We have for* $1 \le m < n \le N$

$$\mathbb{E}v_n^4 \ll n^2 N^{-2\alpha} + n\,N^{-\alpha} \quad \textit{and} \quad \mathbb{E}v_m^2 v_n^2 \ll mnN^{-2\alpha},$$

where the constants implied by $\ll$ *are absolute.*

Before we prove Lemma 8 we finish the proof of Proposition 2. We first apply Theorem B to the martingale $\{V_n^*, \mathcal{F}_n, 1 \le n \le N\}$ with $p = 4$. We obtain

$$\begin{aligned}
\mathbb{E}V_N^{*4} &\le c_4^{-1}\mathbb{E}\left\{\left(\sum_{n\le N} v_n^2\right)^2\right\} \\
&\ll \sum_{n\le N} \mathbb{E}v_n^4 + \sum_{m<n\le N} \mathbb{E}v_m^2 v_n^2 \\
&\ll \sum_{n\le N} n^2 N^{-2\alpha} + \sum_{n\le N} nN^{-\alpha} + \sum_{m<n\le N} mnN^{-2\alpha} \ll N^{4-2\alpha} \ .
\end{aligned}$$

Hence by Lemma 7 and by Theorem A applied to V_N^{*2} we obtain the conclusion of Proposition 2.

Thus it remains to prove Lemma 8. We have

$$\begin{aligned}
\mathbb{E}(v_n^4|\mathcal{F}_{n-1}) &= \mathbb{E}\left\{\left(\sum_{1\le k<n} g(X_n - X_k)\right)^4 |\mathcal{F}_{n-1}\right\} \\
&= \int_0^1 \left(\sum_{1\le k<n} g(u - X_k)\right)^4 du \\
&= \sum_{1\le h,i,j,k<n} \int_0^1 g(u-X_h)g(u-X_i)g(u-X_j)g(u-X_k)du \ .
\end{aligned}$$

By independence and Fubini's theorem the expectation of the integral vanishes unless the 4-tuple (h, i, j, k) consists of two (not necessarily different) pairs of identical numbers. By the proof of Lemma 7,

$$\mathbb{E}\int_0^1 g^2(u - X_k)du = N^{-\alpha} - N^{-2\alpha}$$

and

$$\mathbb{E}\int_0^1 g^4(u - X_k)du \le \mathbb{E}\int_0^1 g^2(u - X_k)du \le N^{-\alpha} \ .$$

Thus

$$\mathbb{E}v_n^4 = \mathbb{E}(\mathbb{E}(v_n^4|\mathcal{F}_{n-1})) \ll n^2 N^{-2\alpha} + nN^{-\alpha} \ ,$$

since there are at most $2n^2$ different pairs and n quadruples.

The proof of the estimate for the mixed moments is similar, but more intricate. We have for $m < n$

$$\mathbb{E}(v_m^2 v_n^2) = \mathbb{E}\left\{ \left(\sum_{1 \le i < m} g(X_m - X_i) \right)^2 \left(\sum_{1 \le k < n} g(X_n - X_k) \right)^2 \right\}$$

$$= \sum_{(*)} \int_0^1 \mathbb{E}\left(g(X_m - X_i) g(X_m - X_j) g(u - X_k) g(u - X_l) \right) du \ ,$$

where the last summation $(*)$ is extended over $1 \le i, j < m$ and $1 \le k, l < n$. We break up this sum into subsums $\Sigma^{(1)}$–$\Sigma^{(5)}$, defined as follows: $\Sigma^{(1)}$ contains all terms with $k, l > m$; $\Sigma^{(2)}$ contains all terms with $k = m$, $l > m$; $\Sigma^{(3)}$ contains all terms with $k = l = m$; $\Sigma^{(4)}$ contains all terms with $k = m$, $l < m$; and $\Sigma^{(5)}$ contains all terms with $k, l < m$. A typical term in $\Sigma^{(1)}$ equals

$$\int_0^1 \mathbb{E}\left(g(v - X_i) g(v - X_j) \right) dv \ \int_0^1 \mathbb{E}\left(g(u - X_k) g(u - X_l) \right) du \ .$$

Such a term will vanish unless both $i = j$ and $k = l$. Thus

$$\Sigma^{(1)} \ll mN^{-\alpha} \cdot nN^{-\alpha} \ll mnN^{-2\alpha} \ .$$

Next, each term in $\Sigma^{(2)}$ vanishes since l is isolated and $\mathbb{E}(g(u - X_l)) = 0$ for fixed u. Similarly, a typical term in $\Sigma^{(3)}$ equals

$$\int_0^1 \int_0^1 \mathbb{E}\left(g(v - X_i) g(v - X_j) \right) g^2(u - v) du dv \ .$$

Such a term will vanish unless $i = j$. In this case the integral equals

$$\int_0^1 \int_0^1 \int_0^1 g^2(v - w) g^2(u - v) du dv dw \ll N^{-2\alpha} \ .$$

Thus

$$\Sigma^{(3)} \ll mN^{-2\alpha} \ .$$

Now a typical term in $\Sigma^{(4)}$ equals

$$\int_0^1 \int_0^1 \mathbb{E}\left(g(v - X_i) g(v - X_j) g(u - v) g(u - X_l) \right) du dv \ .$$

This expectation vanishes unless $i = j = l$. In this case the integral reduces to

$$\int_0^1 \int_0^1 \int_0^1 g^2(v - w) g(u - v) g(u - w) du dv dw \ll N^{-2\alpha},$$

and so

$$\Sigma^{(4)} \ll mN^{-2\alpha} \ .$$

Finally, in $\Sigma^{(5)}$ a typical term equals

$$\int_0^1 \int_0^1 \mathbb{E}\left(g(v - X_i) g(v - X_j) g(u - X_k) g(u - X_l) \right) du dv \ .$$

If one index is different from any of the three other indices, the expectation vanishes. Thus for the expectation not to vanish the indices must come in pairs or as a quadruple. In either case the integral does not exceed $N^{-2\alpha}$. Consequently

$$\Sigma^{(5)} \ll m^2 N^{-2\alpha} \ .$$

This completes the proof of Lemma 8, and thus that of Proposition 2.

This is perhaps an appropriate place to present a rough sketch of a proof of (1.9). The general idea is to keep track of the exponential bounds underlying the proof of Proposition 5.1 in Dehling, Denker and Philipp [6]. This can be done since Lemma 6 above provides the martingale structure needed for the modification of that proof. Proposition 5.1 in the cited paper will then replace Lemma 5 above. Lemma 5.2 in [6] can be dealt with by applying the exponential bound of Kuelbs [17]. For Lemma 5.3 in [6], the argument in Philipp [20, pp. 718–720] can be modified. The proof of Lemma 5.4 in [6] again depends on exponential bounds for martingales, and Lemma 5.5 of [6] is easily modified as $||g||_\infty \leq 1$. At this point there are two options to complete the proof of (1.9). The first option is to modify the proof of Proposition 5 starting with Lemma 2, but replacing the bound in Lemma 5 by the exponential bound mentioned above. The second option is to argue as in Philipp [20, pp. 720–722].

6. Proofs of Theorems 1 and 3

The only difference between U_N and V_N is that in V_N the indicator $\mathbf{1}_L$ has period 1, whereas the indicator in the definition of U_N does not. We shall estimate the relevant difference as follows.

$$\begin{aligned} 0 &\leq \sum_{1\leq i<j\leq N} \left(\mathbf{1}_{[0,t]}(X_j - X_i) - \mathbf{1}\left(0 \leq X_j - X_i \leq t\right)\right) \\ &= \sum_{1\leq i<j\leq N} \left(\mathbf{1}\left(0 \leq X_j - X_i \leq t \mod 1\right) - \mathbf{1}\left(0 \leq X_j - X_i \leq t\right)\right) \\ &= \sum_{1\leq i<j\leq N} \mathbf{1}\left(0 \leq X_j \leq X_i - 1 + t\right) \ . \end{aligned} \tag{6.1}$$

The last equality follows from the fact that

$$0 \leq X_j - X_i \leq t \mod 1$$

if and only if either

$$0 \leq X_j - X_i \leq t, \quad \text{i.e.,} \quad X_i \leq X_j \leq X_i + t$$

or

$$X_i - 1 \leq X_j \leq X_i - 1 + t, \quad \text{i.e.,} \quad 0 \leq X_j \leq X_i - 1 + t \ .$$

The following estimate will yield Theorem 1.

LEMMA 9. *With probability* 1, *we have*

$$\sup_{0\le t\le N^{-\alpha}}\left|\sum_{1\le i<j\le N}\left(\mathbf{1}\left(0\le X_j\le X_i-1+t\right)-\frac{1}{2}t^2\right)\right|\ll N^{\frac{3}{2}(1-\alpha)}(\log N)^4 .$$

Proof. The proof uses the argument of Proposition 1. Let $0\le s\le 1$ and let the binary expansion be given by (2.3). Let $1\le Q\le H$ and define M by (2.4). Then (2.5) holds. For $0\le u<v\le 1$ we define

$$Y(u,v)=Y(u,v;H,Q)=\sum_{H\le j<H+Q}\sum_{1\le i<j}\eta_{ik} ,$$

where we set

$$\eta_{ij}:=\mathbf{1}\left(X_i-1+uH^{-\alpha}\le X_j\le X_i-1+vH^{-\alpha}\right)-\frac{1}{2}(v^2-u^2)H^{-2\alpha} .$$

Then for each s with $0\le s\le 1$ relation (2.6) holds with Z replaced by Y:

$$|Y(0,s)|\le\sum_{m=1}^{M}|Y(d_m2^{-m},(d_m+1)2^{-m})|+2Q(H+Q)2^{-M}H^{-2\alpha} . \tag{6.2}$$

Here the d_m's are integers with $0\le d_m<2^m, 1\le m\le M$. Similarly, relation (2.7) holds with Z replaced by Y. Figure 1 will be useful to establish the

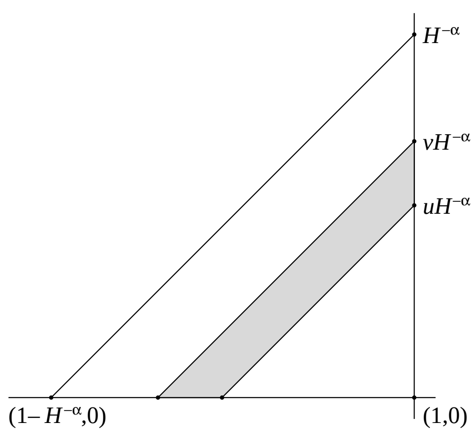

FIGURE 1

following estimates:

$$|\mathbb{E}\eta_{ij}\eta_{kl}|=\begin{cases}=0 & \text{if all four indices are different,}\\ <\frac{1}{2}(v^2-u^2)H^{-2\alpha} & \text{if } i=k,\ j=l,\\ <(v-u)^2H^{-3\alpha} & \text{if } i=k \text{ and } j\ne l \text{ or } i\ne k \text{ and } j=l.\end{cases}$$

The first relation is obvious by independence and since $\mathbb{E}\eta_{ij} = 0$. The second estimate is clear and for the last estimate we note that if $i = k$ and $j \neq l$ the expectation of the product of the indicators figuring in the definition of η_{ij} and η_{kl} equals because of independence

$$\begin{aligned}
&\mathbb{P}(X_i - 1 + uH^{-\alpha} \leq X_j, X_l \leq X_i - 1 + vH^{-\alpha}) \\
&\quad = \int_{1-H^{-\alpha}}^{1} \mathbb{P}(s - 1 + uH^{-\alpha} \leq X_j, X_l \leq s - 1 + vH^{-\alpha}) ds \\
&\quad = \int_{1-H^{-\alpha}}^{1} (\mathbb{P}(s - 1 + uH^{-\alpha} \leq X_j \leq s - 1 + vH^{-\alpha}))^2 ds \\
&\quad < (v-u)^2 H^{-3\alpha} \ .
\end{aligned}$$

Thus for fixed $0 \leq d < 2^m$ and $m \leq M$ we obtain

$$\begin{aligned}
\mathbb{E}(Y(d2^{-m},\, (d+1)2^{-m}))^2 &\leq Q(H+Q)H^{-2\alpha}((d+1)^2 - d^2) \cdot 2^{-2m} \\
&\quad + (H+Q)Q^2 \cdot 2^{-2m} H^{-3\alpha} + (H+Q)^2 Q 2^{-2m} H^{-3\alpha} \\
&\ll H^{1-2\alpha} Q \cdot 2^{-m} + H^{2-3\alpha} Q \cdot 2^{-2m} \ .
\end{aligned}$$

Thus, by (6.2) and (2.7),

$$\begin{aligned}
\mathbb{P}(\sup_{0 \leq s \leq 1} |Y(0,s)| \geq 2R) &\ll \sum_{m=1}^{M} \frac{M^2}{R^2} \cdot 2^m (H^{1-2\alpha} Q \cdot 2^{-m} + H^{2-3\alpha} Q \cdot 2^{-2m}) \\
&\ll R^{-2} H^{1-2\alpha} Q (1 + H^{1-\alpha})(\log H)^3 \ .
\end{aligned}$$

We now can complete the proof of the lemma in the same way as the proof of Proposition 5. □

Combining Lemma 9 with Proposition 1 we get Theorem 1. In view of (6.1) the following rather crude estimate will yield Theorem 3.

Lemma 10. *With probability* 1 *we have*

$$\sum_{1 \leq i < j \leq N} \mathbf{1}\left(0 \leq X_j \leq X_i - 1 + N^{-\alpha}\right) \ll (N^{2-2\alpha} + N^{1+\rho-\alpha})(\log N)^2 \ .$$

Proof. By monotonicity it is enough to prove the lemma for $N = 2^r$, replacing $N^{-\alpha}$ by $2 \cdot 2^{-r\alpha}$. To apply Markov's inequality we need to estimate

$$\mathbb{P}(0 \leq \{n_j \omega\} \leq \{n_i \omega\} - 1 + 2 \cdot 2^{-r\alpha}).$$

Suppose that $i < j$. By a well-known elementary estimate we have for integrable f and $0 \leq a < b \leq 1$ and some $|\theta| \leq 4$

$$\int_a^b f(\{n_j \omega\}) d\omega = (b-a) \int_0^1 f(\omega) d\omega + \theta n_j^{-1} \int_0^1 |f(\omega)| d\omega \ . \tag{6.3}$$

The above probability is bounded by

$$\int_0^1 \mathbf{1}\left(1-2^{-r\alpha+1} \leq \{n_j\omega\}\right) \mathbf{1}\left(\{n_i\omega\} \leq 2^{-r\alpha+1}\right) d\omega \ . \tag{6.4}$$

Now $\mathbf{1}\left(0 \leq \{n_i\omega\} \leq 2^{-r\alpha+1}\right)$ is the indicator of a union of n_i disjoint intervals of length $2^{-r\alpha+1}n_i^{-1}$. Thus using (6.3) n_i times with $f(\omega) = \mathbf{1}(1-2^{-r\alpha+1} \leq \{n_j\omega\})$ and (a,b) the corresponding n_i intervals we obtain for (6.4) the bound

$$\begin{aligned} &\mathbb{P}(1-2^{-r\alpha+1} \leq \{n_j\omega\})\ \mathbb{P}(\{n_i\omega\} \leq 2^{-r\alpha+1}) + 4n_in_j^{-1}\mathbb{P}(1-2^{-r\alpha+1} \leq \{n_j\omega\}) \\ &\leq 2^{-2r\alpha+2} + 4 \cdot 2^{-r\alpha+1} n_i/n_j \ . \end{aligned} \tag{6.5}$$

Now

$$\begin{aligned} \sum_{1\leq i<j<2^r} n_i/n_j &\leq \sum_{1\leq i<j<2^r} (1+cj^{-\rho})^{i-j} \\ &\leq \sum_{j\leq 2^r} (1+cj^{-\rho})(1-1/(1+cj^{-\rho}))^{-1} \\ &\ll \sum_{j\leq 2^r} j^{-\rho} \ll 2^{r(1+\rho)} \ . \end{aligned}$$

Summing (6.5) over $1 \leq i < j \leq 2^r$ we obtain

$$\sum_{1\leq i<j\leq 2^r} \mathbb{P}\left(0 \leq \{n_j\omega\} \leq \{n_i\omega\} - 1 + 2\cdot 2^{-r\alpha}\right) \ll 2^{2r(1-\alpha)} + 2^{r(1+\rho-\alpha)} \ .$$

By Markov's inequality this implies that with probability 1

$$\sum_{1\leq i<j\leq 2^r} \mathbf{1}\left(\{n_j\omega\} \leq \{n_i\omega\} - 1 + 2^{-r\alpha+1}\right) \ll \left(2^{2r(1-\alpha)} + 2^{r(1+\rho-\alpha)}\right) r^2 \ .$$

This proves the lemma for $N = 2^r$, and thus as noted above for N in general. □

Hence by Proposition 3 and (6.1) we have with probability 1

$$\Gamma_N(\alpha) \ll \left(N^{1-\frac{1}{2}\alpha+\frac{3}{2}\rho} + N^{2-2\alpha}\right) (\log N)^{11/2} \ ,$$

which proves Theorem 3.

References

[1] R. C. Baker, *Metric number theory and the large sieve*, J. London Math. Soc. (2) **24** (1981), 34–40.

[2] S. Borovkova, R. Burton, and H. Dehling, *Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation*, Trans. Amer. Math. Soc. **353** (2001), 4261–4318.

[3] D. L. Burkholder, *Distribution function inequalities for martingales*, Ann. Probability **1** (1973), 19–42.

[4] J. W. S. Cassels, *An extension of the law of the iterated logarithm*, Proc. Cambridge Philos. Soc. **47** (1951), 55–64.

[5] K. L. Chung, *An estimate concerning the Kolmogorov limit* distribution, Trans. Amer. Math. Soc. **67** (1949), 36-50.

[6] H. Dehling, M. Denker, and W. Philipp, *Invariance principles for von Mises and U-statistics*, Z. Wahrsch. Verw. Gebiete **67** (1984), 139–167.

[7] ______, *The almost sure invariance principle for the empirical process of U-statistic structure*, Ann. Inst. H. Poincaré Probab. Statist. **23** (1987), 121–134.

[8] M. Drmota and R. F. Tichy, *Sequences, discrepancies, and applications*, Springer-Verlag, Berlin, 1997.

[9] V. R. Eastwood and L. Horváth, *Limit theorems for short distances in* $\mathbf{r}^m$, Statist. Probab. Lett. **45** (1999), 261–268.

[10] W. Eberl and R. Hafner, *Die asymptotische Verteilung von Koinzidenzen*, Z. Wahrsch. Verw. Gebiete **18** (1971), 322–332.

[11] I. S. Gál and J. F. Koksma, *Sur l'ordre de grandeur des fonctions sommables*, Nederl. Akad. Wetensch. Proc. **53** (1950), 638–653; Indagationes Math. **12** (1950), 192–207.

[12] H. S. Green, *The molecular theory of fluids*, North-Holland, Amsterdam, 1952.

[13] K. Hess and C. T. Sah, *Dipole scattering at the* $Si \rightarrow Si0_2$ *interface*, Surface Science **47** (1975), 650–654.

[14] L. Horváth, *Short distances on the line*, Stochastic Process. Appl. **39** (1991), 65–80.

[15] R. A. Hunt, *On the convergence of Fourier series*, Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967), Southern Illinois Univ. Press, Carbondale, Ill., 1968, pp. 235–255.

[16] J. Killingbeck and G. H. A. Cole, *Mathematical techniques and physical applications*, Academic Press, New York, 1971, Pure and Applied Physics, Vol. 35.

[17] J. Kuelbs, *Kolmogorov's law of the iterated logarithm for Banach space valued random variables*, Illinois J. Math. **21** (1977), 784–800.

[18] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, New York, 1974.

[19] H. L. Montgomery, *The pair correlation of zeros of the zeta function*, Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), Amer. Math. Soc., Providence, R.I., 1973, pp. 181–193.

[20] W. Philipp, *Empirical distribution functions and strong approximation theorems for dependent random variables. A problem of Baker in probabilistic number theory*, Trans. Amer. Math. Soc. **345** (1994), 705–727.

[21] Z. Rudnick and P. Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Math. J. **81** (1996), 269–322.

[22] ______, *The pair correlation function of fractional parts of polynomials*, Comm. Math. Phys. **194** (1998), 61–70.

[23] Z. Rudnick and A. Zaharescu, *A metric result on the pair correlation of fractional parts of sequences*, Acta Arith. **89** (1999), 283–293.

[24] B. W. Silverman, *Limit theorems for dissociated random variables*, Advances in Appl. Probability **8** (1976), 806–819.

István Berkes, Mathematical Institute, Hungarian Academy of Sciences, Reáltanoda u. 13–15, H-1053 Budapest, Hungary
E-mail address: berkes@renyi.hu

Walter Philipp, Department of Statistics, University of Illinois, Urbana, IL 61801, USA
E-mail address: philipp@stat.uiuc.edu

Robert Tichy, Institut für Mathematik A, TU Graz, Steyeregasse 30, 8010 Graz, Austria
E-mail address: tichy@weyl.math.tu-graz.ac.at